A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation

Kaufman, C K ; Mosimann, C ; Fan, Z P ; Yang, S ; Thomas, A J ; Ablain, J ; Tan, J L ; Fogley, R D ; van Rooijen, E ; Hagedorn, E J ; Ciarlo, C ; White, R M ; Matos, D A ; Puller, A-C ; Santoriello, C ; Liao, E C ; Young, R A ; Zon, L I

Abstract: The “cancerized field” concept posits that cancer-prone cells in a given tissue share an oncogenic mutation, but only discreet clones within the field initiate tumors. Most benign nevi carry oncogenic BRAFV600E mutations but rarely become melanoma.The zebrafish crestin gene is expressed embryonically in neural crest progenitors (NCPs) and specifically reexpressed in melanoma. Live imaging of transgenic zebrafish crestin reporters shows that within a cancerized field (BRAFV600E-mutant; p53-deficient), a single melanocyte reactivates the NCP state, revealing a fate change at melanoma initiation in this model. NCP transcription factors, including sox10, regulate crestin expression. Forced sox10 overexpression in melanocytes accelerated melanoma formation, which is consistent with activation of NCP genes and super-enhancers leading to melanoma. Our work highlights NCP state reemergence as a key event in melanoma initiation.

DOI: https://doi.org/10.1126/science.aad2197

Originally published at:
Kaufman, C K; Mosimann, C; Fan, Z P; Yang, S; Thomas, A J; Ablain, J; Tan, J L; Fogley, R D; van Rooijen, E; Hagedorn, E J; Ciarlo, C; White, R M; Matos, D A; Puller, A-C; Santoriello, C; Liao, E C; Young, R A; Zon, L I (2016). A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. Science, 351(6272):aad2197.
DOI: https://doi.org/10.1126/science.aad2197

HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2016 May 16.

Published in final edited form as:

Science. 2016 January 29; 351(6272): aad2197. doi:10.1126/science.aad2197.

A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation

Charles K. Kaufman[1,2,3,11], **Christian Mosimann**[4], **Zi Peng Fan**[12,14], **Song Yang**[1], **Andrew Thomas**[1], **Julien Ablain**[1,2], **Justin L. Tan**[15], **Rachel D. Fogley**[1], **Ellen van Rooijen**[1], **Elliott Hagedorn**[1], **Christie Ciarlo**[1], **Richard White**[6], **Dominick Matos**[7], **Ann-Christin Puller**[8], **Cristina Santoriello**[1,9], **Eric Liao**[2,10,11], **Richard A. Young**[12,13], and **Leonard I. Zon**[1,2,3,9,11,*]

[1]Stem Cell Program and Division of Hematology/Oncology, Children's Hospital Boston, Howard Hughes Medical Institute, Boston, MA 02115, USA [2]Harvard Stem Cell Institute, Boston, MA 02115, USA [3]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA [4]Institute of Molecular Life Sciences, University of Zürich, 8057 Zürich, Switzerland [6]Memorial Sloan Kettering Cancer Center, Weill Cornell Medical College, New York, NY 10075, USA [7]Massachusetts General Hospital Cancer Center, Harvard Medical School, Charlestown, Massachusetts 02129, USA [8]Research Institute Children's Cancer Center Hamburg and Clinic of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany [9]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA [10]Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA 02114, USA [11]Harvard Medical School, Boston, MA 02115, USA [12]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA [13]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA [14]Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA 02139 [15]Genome Institute of Singapore, Singapore

Abstract

The "cancerized field" concept posits that cells in a given tissue share an oncogenic mutation or insult and are thus cancer-prone, yet only discreet clones within the field initiate tumors. Nearly all benign nevi carry oncogenic $BRAF^{V600E}$ mutations, but they only rarely become melanoma. The

*Correspondence to: zon@enders.tch.harvard.edu.

Author Contributions:

CKK wrote the manuscript with input from all authors and designed, performed, and interpreted experiments except as noted. CM designed and performed the initial cloning and aided in the characterization of the *crestin* element and variants, and designed experiments. AT contributed to animal husbandry and transgenic injections. JA designed and aided in the interpretation of the CRISPR/Cas9 system used. EH and CC generated embryo videos. RW assisted with microarray and GSEA analysis and experiment design and developed zebrafish melanoma cell and scale culture conditions. JT and RF performed H3K27Ac ChIP-Seq on A375 cells. EVR performed H3K27Ac ChIP-Seq on the primary zebrafish melanoma sample. SY, ZPF, and RY analyzed and interpreted the ChIP-Seq and ATAC-Seq data. DM helped develop the scale transplant protocol. ACP subcloned *crestin_1kb:EGFP*. CS provided embryo *in situ's* for *crestin*. EL provided the *sox10:mCh* line. LIZ helped conceive the study and interpreted experiments.

Supplementary Materials

Figs. S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12

Movie S1, S2, S3

Table S1-S6 (Microarray Data Annotated, Super-Enhancers A375, Primers, RNA-Seq crestin + cells, RNA-Seq human melanoma, GSEA Gene Lists)

zebrafish *crestin* gene is expressed embryonically in neural crest progenitors (NCP's) and is specifically re-expressed in melanoma. We show by live imaging of transgenic zebrafish *crestin* reporters that, within a cancerized field (*BRAF*^{V600E}-mutant; *p53*-deficient), a single melanocyte reactivates the NCP state, and this establishes that a fate change occurs at melanoma initiation in this model. We show the *crestin* element is regulated by NCP transcription factors, including *sox10*. Forced *sox10* overexpression in melanocytes accelerated melanoma formation, consistent with activation of a NCP gene signature and super-enhancers leading to melanoma. Our work highlights the importance of NCP state reemergence as a key event in melanoma initiation.

Introduction

While the key importance of oncogene activation and tumor suppressor inactivation in tumor formation is well appreciated, our understanding of the early events of cancer initiation remains limited. The mechanisms that enable only sporadic cells to complete the conversion to a malignant state amongst a large group of cancer-prone cells, sometimes described as a "cancerized field," remain unclear (1). Better characterizing initiating events would identify targets for early therapeutic interventions and also provide prognostic information about which pre-cancerous lesions are most worrisome for progressing. Melanoma is a cancer of transformed melanocytes, which are pigment-producing cells derived from the embryonic neural crest lineage, and is frequently driven by *BRAF* or *RAS* mutations (~80% of case) (2, 3). Melanoma is treatable and curable when it is localized and can be resected completely, but remains largely incurable once it has spread, even when treated with new kinase- and immune checkpoint-targeted therapies (4). Our lab previously developed the first animal model of a *BRAF*^{V600E}-driven cancer by placing the human *BRAF*^{V600E} gene under the control of the melanocyte-specific *mitfa*-promoter in transgenic zebrafish (5). When crossed into a *p53* mutant loss-of-function background, these zebrafish (referred to here as *p53/BRAF)* invariably develop nevi and, after several months, invasive melanoma (5). Despite creating this extensive "cancerized field" in which all melanocytes harbor both oncogenic *BRAF*^{V600E} and *p53* loss throughout their lifespan, these *p53/BRAF* melanoma-prone zebrafish typically develop one to three melanoma tumors after several months of age, indicating that other molecular alterations are important for tumor initiation.

crestin transgenics mark neural crest

To investigate the dynamics and mechanism of the observed sporadic melanoma formation, we aimed to visualize and characterize melanoma lesions at the time of their initiation. The functionally uncharacterized zebrafish *crestin* gene marks the neural crest during embryonic development and then becomes undetectable by ~72 hours post fertilization (hpf) (6, 7), but we previously found that it specifically re-expresses in melanoma tumors in adult zebrafish (8). We reasoned that a *crestin*-based reporter transgene would allow us to track embryonic neural crest cells as well as melanoma tumors *in vivo*, potentially from their earliest onset. We PCR-amplified a 4.5 kb upstream region common to multiple *crestin* insertions in the zebrafish genome, and cloned this element upstream of an *EGFP* reporter (Fig 1A, *crestin:EGFP*). In stable transgenic zebrafish embryos, this construct reproduced *crestin* mRNA expression by EGFP fluorescence (Fig 1B, C, S1A), and time-lapse videos demonstrated the dorsal emergence and wide migration of these *crestin*-expressing putative

neural crest progenitor cells (Movie S1, S2). Neural crest expression was reproducible in multiple independent lines and with additional reporter genes (see below, *creERT2* and *mCherry*, S5A–C). As with endogenous *crestin* expression, transgenic *crestin:EGFP* expression was not detectable after 3 days post-fertilization (dpf) and did not come back on in wild type juvenile or adult zebrafish.

To confirm that the *crestin* transgenes target neural crest progenitors, we also generated transgenics for *crestin:creERT2* to genetically mark *crestin* expressing embryonic cells using a Cre/lox-dependent *EGFP*-to-*mCherry* switching line (9) and genetically labeled neural crest-derived cells including melanocytes/pigment cells (red cells in Fig 1D, E), jaw cartilage (Fig 1F), and lateral line glia (Fig 1G). As the *crestin* gene is specific to zebrafish, we wanted to ensure that *crestin* reporter embryonic expression is consistent with another conserved early neural crest marker, the transcription factor *sox10*. Confocal analysis of double-transgenic *Tg(crestin_1kb:EGFP)* and *Tg(sox10:mCh)* (10) zebrafish embryos showed a high degree of overlap in reporter gene expression (Fig 1H) with any differences matching published *in situ* hybridization (ISH) data (11). Thus, our *crestin* transgenic lines recapitulate *crestin* expression and specifically mark the embryonic neural crest stem/progenitor cell population.

crestin transgenics visualize melanoma initiation

We next determined if *crestin:EGFP* is re-expressed in melanoma tumors as noted previously by ISH (8). We found *crestin:EGFP* is expressed in tumors arising on triple transgenic *p53/BRAF/crestin:EGFP* adult zebrafish but is absent in the remainder of the animal, highlighting its specificity to the tumor (Fig 2A). We next followed developing zebrafish to observe the onset of *crestin:EGFP* (+) expression. We found *EGFP*-expressing patches of cells that had not yet formed raised melanoma lesions (Fig 2B) and even detected single isolated *EGFP* (+) cells in *p53/BRAF/crestin:EGFP* zebrafish (Fig 2C). Although rare events, we could track the persistence and enlargement of single EGFP + cells (Fig S2A, B). Small patches of *EGFP +* cells, containing < 50 cells, are also readily tractable as they enlarge (Fig S2C). Analysis of single scales with discreet *crestin:EGFP* (+) patches demonstrated that transgene expression detectable by fluorescence microscopy overlaps with *crestin mRNA* detected by ISH (Fig 2D). Together, these observations reveal that, after pan-neural crest expression confined to the embryo, our *crestin* reporter expresses specifically and reproducibly in melanoma tumors, thus providing an *in vivo* genetic label for melanoma cells that is significantly earlier that previous detection methods.

We next addressed the dynamics of reemerging *crestin* expression in cohorts of *p53/BRAF/crestin:EGFP* zebrafish. At the population level, *crestin:EGFP* (+) patches of cells (S1B, C) were visible prior to the appearance of grossly raised melanoma lesions (Fig 2E, S1D, Movie S3). The *crestin:EGFP* expression is undetectable in the *p53/BRAF* fish from 3 dpf to > 21 dpf, again consistent with previous *in situ* analyses for endogenous *crestin*. We tracked individual small patches of *crestin:EGFP* (+) cells over time as they progressed into fully formed raised melanoma lesions (Fig 2E), and we found that all melanomas tracked in this manner initiated from *crestin:EGFP* (+) patches of cells (30 out of 30). Thus, if a patch is seen in the *p53/BRAF* background, it will become an overt melanoma. These data

demonstrate that reemergence of *crestin:EGFP* expression, and a neural crest progenitor state, correlates with melanoma initiation in an *in vivo* model of *de novo* melanoma formation.

We then wanted to establish that pre-tumor patches of *crestin:EGFP* (+) cells are tumorigenic and can autonomously expand locally after transplant, so we performed scale auto-transplants on *p53/BRAF/crestin:EGFP* zebrafish (12). Patches of *crestin:EGFP* (+) cells survive following transplant, expand, and the *EGFP* (+) cells persist and further expand when later re-removing the transplanted scale, suggesting the cells have invaded the hypodermis (representative example, Fig 2G). We achieved similar more short-term results with isolated scales placed in tissue culture (Fig S3A, S4A, B) and with allo-transplants to sublethally irradiated recipient zebrafish (Fig S3B) (8). Thus, early patches of *crestin:EGFP*-expressing cells are transplantable in a manner suggesting they are already tumorigenic.

Transcriptional regulators of *crestin* expression

As the *crestin* element proved to be a highly specific and unique tool for monitoring neural crest and melanoma development, we aimed to identify i) a minimal element within the 4.5 kb *crestin* promoter/enhancer that could drive this expression pattern and ii) key transcriptional regulators within the element. Sequence analysis of the *crestin* locus, which is replicated throughout the zebrafish genome >40 times, is similar to another retroelement called *bhikari* that is expressed in early mesendoderm (6, 7, 13) (Fig S6A). Both a 1 kb segment from the putative retroelement promoter region and a smaller 296 bp subregion fully reproduced the neural crest- and melanoma-specfic expression pattern of the full 4.5 kb *crestin* element (Fig 1H, Fig S5 A–G), with slightly weaker expression for the 296 bp element. We conclude that key neural crest regulatory elements are contained in this 296 bp of DNA, although additional contributory binding sites may also be functional in the context of the larger *crestin* element.

Database searches identified multiple predicted transcription factor binding sites for important neural crest developmental regulators within the 296 bp segment, including two *sox10*, one *pax3*, one *E-box (myc* or *mitf*-binding site), and one *tfap2* site (14) (Fig 3A, Table S1). To determine which sites are functionally required for *crestin* transgene expression, we individually mutated the core consensus for each site (Fig S6B, Fig 3A), and tested expression at 24 hpf in F0 embryos injected with the different *crestin_296bp:EGFP* constructs at the 1-cell stage, examining > 80 successfully injected F0 embryos per construct. Whereas mutation of the predicted *pax3* site left the expression pattern largely unchanged from wild type (~55% of embryos with neural crest predominant expression in both), mutation of either *sox10* site drastically reduced neural crest expression (≤20%), as did mutation of the *tfap2* site or the E-box site (< 10% and none, respectively) (Fig 3A, Fig S6C–D). These functional transcription factor binding sites provide an explanation for the neural crest specificity of *crestin* transgene expression, which integrates regulatory signals of multiple neural crest transcription factors including *sox10*.

Neural crest signature in melanoma initiation

To test whether early precursor melanoma lesions express other melanoma and neural crest progenitor markers in addition to *crestin*, we isolated individual scales from *p53/BRAF/Na/MiniCoopR/crestin:EGFP* zebrafish with early *crestin:EGFP* (+) patches and compared them to adjacent individual scales without *crestin:EGFP* expression (Fig 3 B–C) and performed Affymetrix microarrays (15). Such scales appear well-matched in their cell makeup, particularly in regards to melanocytes, as shown by *mitf:mCh* co-expression (marking melanocytes), and *crestin:EGFP* (Fig S7). *Crestin:EGFP* (+) scale-enriched genes include neural crest- (*e.g. crestin, mitf, dlx2a)* and melanoma-expressed (*e.g. mia, mt* [metallothionein]) genes (Table S2) (16–18). We confirmed enrichment by quantitative RT-PCR on independent *crestin:EGFP* (+) and (−) scale (including *crestin, dlx2a,* and *mia*, Fig 3D). We also found *sox10* expression enriched in the *crestin:EGFP* (+) samples (Fig 3D). S*ox10* is a known marker and key regulator of neural crest identity (Figure 1H) (19) and, from our data, of *crestin* expression (Fig 3A); it is also sufficient to direct reprogramming of human fibroblasts to induced neural crest cells (20) and can be highly expressed in melanoma where it is involved in growth control (21–23). We used GSEA analysis to query a rank-ordered list of the *crestin:EGFP* (+) scale enriched genes for an association with all neural crest-expressed genes in the ZFIN database, and we found a significant correlation (FDR q-value and FWER p-Value of 0.019) (Fig 3E, left panel, Table S6). Similarly, we used GSEA analysis to compare a rank-ordered list of genes enriched in ES-derived human neural crest cells (24) with genes enriched ≥2 fold in *crestin:EGFP* (+) scales, and we found a positive correlation, detectable even across species (FDR q-value and FWER p-value of 0.089) (Fig 3E, right panel, Table S6). These data collectively support the concept that key aspects of NCP state reemerge at the time of melanoma initiation, as read out by the *crestin:EGFP* reporter.

Neural crest progenitor identity and melanomagenesis

Based on our analysis of *crestin* expression, which provides an *in vivo* read-out of NCP identity at the time of melanoma initiation, we reasoned that favoring entry into or inappropriately maintaining the NCP state in a cancerized field of melanocytes would accelerate the onset of melanoma formation (Fig 3F). The neural crest master transcription factor *sox10* has been shown to increase *crestin* mRNA in embryos when overexpressed (25), and we found similar results for our *crestin:EGFP* reporter (Fig S8). Misexpression of *SOX10* in postnatal fibroblasts also generates multipotent neural crest cells in culture (20). We therefore forced overexpression of *sox10* in melanocytes using the transgenic *MiniCoopR* system we previously developed (15) and found that the forced overexpression of *sox10* in melanocytes accelerated melanoma onset significantly versus controls (Fig 3G). To examine the consequence of *sox10* inactivation in potentially in inhibiting melanoma formation in our system, we used a melanocyte-specific CRISPR/Cas9 system to target *sox10* in the *p53/BRAF/Na* background (Fig S9A). As compared to controls in which *p53* is redundantly targeted (i.e. already mutated in our system) using an analogous vector, we found a significant slowing of median tumor onset in the *sox10* CRISPR/Cas9 setting [133 days (*p53*) vs 180 days (*sox10*), $p < 0.0001$, Fig S9B). When melanomas developed in the *sox10*-targeted background, the *sox10* target genomic locus exhibited a propensity for mutations that preserve predicted *sox10* function (e.g. point mutations or in-frame deletions,

~60% of sequenced genomes) as opposed to inactivating mutations (e.g. frame-shifts, ~40% of sequenced genomes), suggesting a selective pressure for retention of *sox10* function (Fig S9C–D). These gain- and loss-of-function results together strongly support our hypothesis that reemergence of NCP state is an important event in melanoma tumor initiation.

Neural crest super-enhancers and melanoma

Finally, we aimed to understand how the expression of neural crest genes such as *sox10* may be regulated in zebrafish and human melanoma using a combination of ChIP-Seq and ATAC-Seq. Chromatin regions with high levels of H3K27Ac histone marking have been referred to as super-enhancers, or stretch-enhancers (26, 27), and have been identified as key transcriptional regulatory elements that modulate cell-type specific and cancer-related gene expression (26–29). We used ChIP-seq to identify H3K27Ac enriched regions in a zebrafish *crestin:EGFP* (+) melanoma cell line (zcrest1) that we derived from *p53/BRAF/Na/MiniCoopR/crestin:EGFP* zebrafish and noted significant regions of H3K27Ac enrichment at *crestin* loci, identified as super-enhancers (SE's) (representative locus shown in Fig 4A, SE = red bar). We also identified Sox10 binding by ChIP-Seq across the *crestin* locus (Fig 4A, bottom track), consistent with our promoter analysis linking *sox10* to *crestin* transcriptional regulation (Fig 3A). We examined the *sox10* locus in the zcrest 1 zebrafish melanoma cell line and also identified H3K27Ac SE marks (Fig 4B). These SE's were similarly found at *sox10* and *crestin* via H3K27Ac ChIP-Seq performed on a freshly isolated *primary* zebrafish melanoma tumor (Fig S10A–B, red bars), supporting our findings on the cell lines as being representative of the *in vivo* landscape. ATAC-Seq identified open and accessible chromatin corresponding to the SE's at *crestin* and *sox10* (Fig 4A–B) and other SE-associated loci in two zebrafish melanoma cell lines (Fig S10E–F, see below) (30). These data suggest a molecular basis for the epigenetic state read out by *crestin* of NCP identity in initiating melanoma cells.

We then wanted to compare our fish studies to human melanoma, for which super-enhancer analysis is limited. Examining the Cancer Cell Line Encyclopedia (CCLE) database (31), we found that most human melanoma lines (51/60) express *SOX10* based on Affymetrix microarray data (Fig S11A). As with zebrafish melanomas, ChIP-Seq showed enriched H3K27Ac marks near the *SOX10* locus in six *SOX10*-expressing human melanoma lines tested but not in a rare *SOX10*-negative human melanoma cell line (LOXIMVI, Fig 4C, E, Fig S11A). The *SOX10* SE's were ranked 3 and 6 out of 842 super-enhancers from ~15,000 total enhancers in the A375 line (Fig S11B, Table S3). Clustering based on the SE landscape yielded two distinct groups of *SOX10*-expressing lines that correlated with the presence or absence of expression of melanocyte differentiation markers, *TYR* and *DCT* (Fig S11C–D). H3K4me1, a histone modification typically at active enhancers, was also enriched at *SOX10* in the representative A375 melanoma line (Fig 4C). Remarkably, these *SOX10* SE peaks were also found in published H3K27Ac data from human ES-derived neural crest cells (hNCC's) (32) (Fig 4C, E). Examining multiple normal and cancer cell types (66 and 18 types, respectively), the enrichment of H3K27Ac signal at *SOX10* was significant and specific to melanoma cells, hNCC's, and brain tissue, which contains *SOX10*-expressing oligodendrocytes (Fig 4E) (26). Beyond *SOX10*, a similar SE epigenetic signature was shared for the neural crest transcription factor *DLX2* amongst melanomas across species and

was enriched in melanomas and hNCC's (Fig 4D–E, Fig S10C, E). *DLX2* expression is enriched in sorted *crestin* (+) embryonic neural crest cells (Table S4), in *crestin (+)* precursor melanoma patches (Fig 3D), and in the less differentiated, *TYR/DCT (−)* melanomas relative to cultured normal human melanocytes (Fig S11C–D, F, Table S5). SE's were also found at *TFAP2* family members, whose functional binding site was identified in the 296bp *crestin* element (Fig 3A), in A375 human melanoma (*TFAP2C,* Fig S12A), and in the zebrafish melanoma cell line and primary melanoma (*tfap2a,* Fig S10D, F). Such SE's are also found near other *crestin* (+) cell-enriched genes *MIA* and *MT2A* (metallothionein genes) in A375 melanoma cells (Fig S12B,C).

Discussion

Together, these data corroborate and broaden our conclusions that melanoma precursor cells re-initiate an embryonic neural crest signature and activate a melanoma gene program at their initiation. While the expression of stem cell factors or embryonic genes has been noted previously in advanced malignancies including melanoma, it remained uncertain whether this was simply due to aberrant misexpression of these genes, or whether these genes were present from tumor outset (Fig 3F). Our data support a model in which the stem/progenitor cell gene programs are an integral part of cancer initiation and not re-acquired later (Fig 3F). We note a remarkable conceptual similarity between the reprogramming iPS process and the tumorigenic reemergence of NCP state seen in our system as melanocytes with the required, but insufficient, *p53/BRAF* genetic mutations stochastically reenter a NCP state (as read out by *crestin* expression) early at the initiation of melanoma. As the earliest, initiating stages of more cancers are analyzed, stem cell/progenitor phenotype reacquisition may be a generally observed phenomenon in most cancers (33).

Our work establishes a method to live image cancer development when a tumor starts, potentially as a single cell, providing a unique view of the initiating events. Our model is that NCP identity arises early in the life of a tumor, and our data support this being a necessary step during the early life of the tumor. In our *p53/BRAF/crestin:EGFP* model, all *EGFP* (+) patches of *crestin:EGFP* cells that we track go on to enlarge and form tumors; given the rarity of capturing single cells in this background, we can not rule out that clones may rarely fail to progress to patches. The *crestin* gene has the rare characteristic of being expressed in neural crest cells in embryos, but not being expressed in adult tissues except when melanoma arises. Super-enhancers shared between melanoma and neural crest are specifically activated. The neural crest signature encompasses a combinatorial code including *SOX10, DLX2*, and the *TFAP2* family. It is unlikely that any single of these genes is sufficient for the reprogramming event, but it is the combination of multiple transcription factors that participate. In contrast to *SOX10*, *DLX2* expression is characteristic of genes that are expressed by neural crest and melanoma, but not expressed in melanocytes. As such, this gene set may have diagnostic importance in the initiating cancer cell. Interestingly, a reevaluation of published expression data for nevi, primary melanoma, and metastatic melanoma does reveal increased *SOX10* expression in the malignant melanomas (34). Nevi show a wide range and consistently lower amount of *SOX10* expression than melanoma, raising the possibility that the histologically defined category of nevus is capturing a range

of melanoma initiating capacity (i.e. those nevi expressing higher *SOX10* levels may have initiated or may be more prone to initiating melanoma).

Several major questions remain about an initiating cancer cell that may now be more accessible with our live visualization tool. The niche environment must participate in the process of initiation, perhaps through the activation of neural crest signaling pathways akin to the development of the normal neural crest or by stress pathways related to irradiation or oxidative damage. Further genetic mutations, as opposed to isolated epigenetic changes, may be required for tumor initiation from the cancerized field; however, the absence of identifiable functionally relevant exomic mutations in a study of fifty-three zebrafish melanoma tumors to which we contributed would tend to favor that the key genetic drivers (i.e. $BRAF^{V600E}$ and mutant *p53*) are already present in our model (35). Work on this question will provide information on how cancer initiates. The reprogramming event appears to occur in one melanocyte or progenitor in a cancerized field, and defining why the process initiates in that single cell rather than an adjacent cell will provide an understanding of protective mechanisms in cancer formation.

Materials and Methods

Cloning of *crestin* promoter/enhancer

The transcript for *crestin*, as originally described as AF195881, was used to BLAST the zebrafish genome and, as has been previously noted, identified many partial or complete highly similar (>90% identical) sequences spread throughout the genome (6, 7). Most notably, multiple insertions are present on Chromosome 4, and a ~4.5 kb sequence located upstream of the predicted *crestin* ORF was noted to be present in multiple instances. Reasoning that this segment may contain the relevant regulatory elements of *crestin,* primers were designed to PCR amplify upstream sequences of the *crestin* locus (LOC796814) on Chromosome 4 in the TU background (primers 299 and 302, see Primers Supplemental Table). The 1 kb fragment and 296 bp fragment were isolated using PCR primers 302/340 and 517/516, respectively. Fragments were cloned into pENTR5' (Life Technologies) per manufacturers instructions. Expression vectors were derived using Multisite Gateway® technology per manufacturers specifications and the Tol2 Kit (36) (*EGFP*, vector 383; *mCh*, vector 386; SV40 poly A, vector 302; *EGFP* with mouse minimal β-globin promoter (37); destination vectors, vector 394 alone and modified with addition of *crystallin:YFP* marker; *zmitfa* middle entry vector, gift of Craig Ceol; MiniCoopR from (15)).

Production of transgenic zebrafish and lineage tracing

One cell stage embryos from the AB strain grown under standard, IACUC-approved conditions were injected with the given DNA construct at 25 ng/ul with Tol2 mRNA at 20 ng/ul (36). Embryos were screened at 24 hpf for neural crest expression of *EGFP* or *mCh*, or in the case of *Tg(crestin:CreERt2;crystallin:YFP)*, screened for YFP-positive lenses at 4 dpf. These were grown to adulthood and outcrossed to identify founders that gave germline transmission. For each DNA construct, ≥2 independent lines were generated to confirm the expression pattern. For lineage tracing of *crestin*-expressing cells, a stable transgenic line of *Tg(crestin:creERt2;crystalline:YFP)* was crossed to the *Tg(−3.5ubi:loxP-GFP-loxP-*

mCherry) line (*ubi:Switch*) (9); embryos were collected and treated with 10 μM 4-OHT at 50% epiboly and 24 hpf while grown at 28°C in E3 medium as per standard protocol.

Embryo and adult imaging

Transmitted light and fluorescence images of adult and non-confocal images of embryos were collected on a Zeiss Discovery V.8 Stereoscope with an Axiocam HRc. Static confocal images were collected on a Nikon C2si Laser Scanning Confocal using 25X objective on embryos mounted in 1% low melt agarose. Maximum intensity projections of Z stacks or 3D reconstructions are presented here. Movies of developing embryos were collected on a Nikon Eclipse Ti Spinning Disk Confocal with a 10X objective, with tiled images collected every 6–7 minutes. Images were processed using Photoshop, ImageJ, or Imaris. Multiple tiled images of adult zebrafish were stitched together using the automated Photomerge function in Photoshop or manually aligned.

Melanoma model and *MiniCoopR* system

Experiments were performed as outlined in (15). Briefly, *p53/BRAF/Na* embryos were injected with equal amounts of *MiniCoopR* alone or *MiniCoopR;mitf:sox10* and selected for melanocyte rescue at 48 hours. Equal numbers of melanocyte-rescued embryos were grown to adulthood (n=14 for control and n=13 for mitf:sox10) and scored for the emergence of raised melanoma lesions as per (15). Survival curves and statistics were generated in Prism.

qPCR, Microarray, and *in situ* hybridization

Adult *p53/BRAF/Na/MiniCoopR/crestin:EGFP* fish were anesthetized with Tricaine, viewed under the fluorescent dissecting scope, and precursor patches of *crestin:EGFP* (+) cells were identified and associated single scales removed. These fish contained stable (i.e germ-line transmitted) alleles of *crestin:EGFP* and control *MiniCoopR* (i.e. no test gene). Neighboring *crestin:EGFP* negative scales from the same zebrafish were selected for controls. Scales were immediately placed in Trizol, and total RNA was purified following the manufacturer's instructions with the additional use of Sigma GenElute LPA carrier. RNA was analyzed on a BioAnalyzer, and high quality samples were chosen for microarray libraries generated using the Ovation Pico WTA System V2 and Encore Biotin labeling system for hybridization on the Zebrafish Gene 1.0 ST Affymetrix Array or used for cDNA preparation using SuperScript III. RMA normalized Affymetrix Array results were sorted for maximum fold (log2) increase in gene expression in GFP positive samples versus negative samples. qPCR reactions were run on triplicate biological samples with triplicate technical replicates and normalized to β-actin expression, with representative results presented. For *in situ* hybridizations, we used the methods of (38) and the *crestin* probe from (7).

ChIP-seq and super-enhancer analysis

Human melanoma cells were grown to confluence in DMEM + 10% FCS, and ~ $1X10^8$ cells were formaldehyde crosslinked and collected. ChIP was performed using the methods of (30, 39) with antibodies against H3K27Ac (ab4729) (Abcam), H3K4Me1 (ab8895) (Abcam), and SOX10 (Santa Cruz Biotechnology, sc-17342x). Libraries were prepared using the NEBNext Multiplex Oligos for Illumina kit (NEB) and run on an Illumina Hi-

Seq2000. Data analysis, including enhancer and super-enhancer calling, were performed as described in (26). Genomic track images were generated using the IGV package (40) and the UCSC Browser (41).

All human ChIP-Seq datasets were aligned to build version NCBI37/HG19 of the human genome using Bowtie (version 0.12.9) (42) with the following parameters: –n2, -e70, -m2, -k2, --best. We used the MACS version 1.4.1 (Model based analysis of ChIP-Seq) (43) peak finding algorithm to identify regions of ChIP-Seq enrichment over background. A p-value threshold of enrichment of 1e-9 was used for all datasets. Wiggle files for gene tracks were created using MACS with options –w –S –space=50 to count reads in 50bp bins. They were normalized to the total number (in millions) of mapped reads producing the final tracks in units of reads per million mapped reads per bp (rpm/bp).

Identifying super-enhancers

The identification of super-enhancers has previously been described in detail (26). Briefly, H3K27ac peaks were used to identify constituent enhancers. These were stitched if within 12.5kb, and peaks fully contained within +/− 2kb from a TSS were excluded from stitching. H3K27ac signal (less input control) was used to rank enhancers by their enrichment. Super-enhancers were assigned to active genes using the ROSE software package (younglab.wi.mit.edu/super_enhancer_code.html).

ATAC-Seq

Zebrafish melanoma cell lines were grown to 80% confluence, trypsinized, counted, and 50,000 cells were lysed and subjected to “tagmentation” reaction and library construction as described in (30). Libraries were run on an Illumina Hi-Seq2000. All zebrafish ATAC-Seq datasets were aligned to build version Zv9 of the zebrafish genome using Bowtie2 (version 2.2.1) (42) with the following parameters: --end-to-end, -N0, -L20. We used the MACS2 version 2.1.0 (43) peak finding algorithm to identify regions of ATAC-Seq peaks, with the following parameter --nomodel --shift −100 --extsize 200. A q-value threshold of enrichment of 0.05 was used for all datasets.

Scale transplants

Adult *p53/BRAF/Na/MiniCoopR/crestin:EGFP* were anesthetized, viewed under the fluorescent dissecting scope, and precursor patches of *crestin:EGFP* (+) cells were identified (not from raised melanoma lesions) and single associated scales were removed and placed in 50 μl drops of E3 buffer on a petri dish lid. Anesthetized recipient zebrafish (in the case of allotransplants) or the same zebrafish (in the case of autotransplants) were gently placed on a wet sponge and a recipient site selected (free of *crestin:EGFP*-expressing cells). One scale was removed at the donor site, and the previously selected donor scale from the drop of E3 was placed on the zebrafish and slid posterior to anterior into place using surrounding scales to hold it in place. Recipients were quickly placed in fresh zebrafish water and monitored for recovery from anesthesia. Transplants were monitored frequently with some loss of transplanted scales (~20%) occurring quickly within one day due to simple dislodgement. Once in place for ~ 4 days, scales were firmly incorporated and could be monitored over time and photographed at the same magnification under the dissecting scope following mild

Tricaine anesthesia of the recipient zebrafish. In single scale autotransplants in a representative cohort of 10 fish, 3 scales were lost in the first week (we consider a technical failure with the scales falling out), 5 showed expansion of the *crestin:EGFP* cells, and 3 showed no change or loss of *crestin*:EGFP cells. In single scale allotransplants onto sublethally irradiated *casper* recipient fish, 8 *EGFP +* scales were transplanted onto different recipients with 1 scale lost in the first week, 5 scales showing expansion of the *crestin:EGFP* cells for ≥2.5 weeks, and 2 showing stable appearance for ≥2 weeks. 6 of 6 *EGFP (−)* scales showed loss of pigmented cells in this time frame.

Promoter Analysis

Transcription factor binding sites were predicted using JASPAR (14). We used the Q5® Site-Directed Mutagenesis Kit from NEB to introduce mutations to destroy the chosen transcription factor binding site (see Primers Supplemental Table). To analyze expression *in vivo,* equal volumes of a mixture of the *crestin_296bp:EGFP* construct variant at 20 ng/ul mixed with 5 ng/ul of *ubi:mCh* (9, 36) (for an injection control) were injected with 20 ng/ul of Tol2 mRNA into single cell AB embryos. At 24 hpf, embryos were fixed in PFA, washed and stored in PBS, and scored for *mCh* expression to identify successfully injected embryos and for *EGFP* to bin based on the predominant expression pattern. More than 80 transgenic F0 embryos for each construct were scored.

Investigating SOX10 and DLX2 super-enhancers in healthy normal cells and melanoma cells

To investigate whether the two SOX10-asscociated super-enhancers identified in A375 cells are also present in other cell types, they were compared to the super-enhancers identified in other melanoma cell types, neural crest cells, and 84 additional cell types from normal or cancer cells described in (26).

First, the SOX10-asscociated super-enhancers were operationally called “present” in a cell type if the super-enhancers identified in the cell type overlap with the SOX10 super-enhancers in A375 cells by at least 1bp. Second, the H3K27ac signal density of the SOX10-asscociated super-enhancers identified in A375 cells was compared to the mean H3K27ac signal density of all enhancer clusters in each cell type. For each cell type, the average H3K27ac ChIP-seq read density was calculated in reads per million mapped reads per bp (r.p.m./bp) for the two SOX10 super-enhancers identified in A375 cells as well as all enhancer clusters identified using the ROSE software package in the cell type (younglab.wi.mit.edu/super_enhancer_code.html). The fold difference of H3K27ac ChIP-seq signal at SOX10 super-enhancers over the mean H3K27ac ChIP-seq signal at all enhancer in each cell type was plotted in Fig 4E. The same process was undertaken for the SE’s at the *DLX2* locus.

Pairwise comparison of super-enhancers between different melanoma cell lines and neural crest cells

The set of super-enhancer regions in each cell type (SKMEL2, SKMEL30, UACC257, LOXIMVI, A375, CJM, COLO679, and neural crest cells) were merged together if overlapping by 1 bp, resulting a total of 3407 merged super-enhancer regions. The neural

crest cell data was previously published by (32). The average H3K27ac ChIP-seq read density was calculated in reads per million mapped reads per bp (r.p.m./bp) for each of the merged regions. The pair-wise comparisons by Pearson correlation were performed on all data sets using the average read density at the merged regions. The average linkage hierarchical clustering of the Pearson correlation was shown in the heatmap (Fig S11C).

Zebrafish melanoma cell lines and *in vitro* scale imaging

A single EGFP (+) melanoma tumor arising in a *p53/BRAF/Na* fish injected with *MiniCoopR* plasmid and *crestin:EGFP* plasmid (zcrest 1 line) or a *p53/BRAF/Na/crestin:EGFP* fish injected with MiniCoopR plasmid (zcrest 2 line) was removed after sacrificing the adult fish. Briefly, tumors were dissociated with a razor blade and trypsin, filtered, and plated on a fibronectin coated well and grown in rich media supplemented with FBS and zebrafish embryo extract as described (44). After several passages, the zcrest 1 line was sorted for EGFP + cells, which were continued as the line. For the zcrest 2 line, after several passages, most if not all cells remaining were *EGFP*+. Both lines continue to be EGFP (+) and have been grown for > 50 passages on plastic in standard DMEM + 10% FBS with 1X GlutaMAX™ supplement and penicillin/streptomycin antibiotics.

Scales from *p53/BRAF/Na/crestin:EGFP/MiniCoopR* and *p53/BRAF/crestin:EGFP* fish with and without *EGFP*+ patches of cells were placed in zebrafish melanoma growth medium in fibronectin coated wells in 384-well format, flat-bottomed plates and imaged daily on a Yokogawa CV7000 confocal imager with brightfield and z-stack image projections collected.

GSEA Analysis

Using our microarrays comparing *crestin:EGFP* (+) *vs* (−) scale gene expression, zebrafish genes were rank-ordered (10,705 genes) from high to low for enrichment (crestin scales rank.rnk). We generated a gene list of all neural-crest expressed genes in the ZFIN database (317 genes, zfincomp.gmt) (45), and used the preranked GSEA analysis tool. Using a list of genes with ≥2 fold enrichment in *crestin:EGFP* (+) scales by microarray (Crestin Scales.gmt), we queried a list of 17,575 genes rank-ordered for their enrichment in human neural crest cells versus parental ES cells from (32) (Rada Ranked.rnk) also using the preranked GSEA analysis tool (46).

RNAseq Analysis

Multiple human melanoma cell lines and adult human epidermal melanocytes (purchased from Life Technologies) were grown to near confluence, and total RNA isolated using the standard Trizol protocol. Illumina libraries were prepared using Ribo-Zero Magnetic Gold Kit (epicenter) and NEBNext Ultra RNA Library Prep Kit (NEB) and run on a Hiseq 2500, reads aligned using Tophat 2.0, and FPKM values determined using Cufflinks. For sorted *crestin* (+) cells, transgenic *crestin_1kb:EGFP* adults were mated, and embryos collected and grown to the 15 somite stage. These were homogenized, filtered, and sorted using FACS into PBS, collected ~5,500 *EGFP* (+) cells and 100K EGFP (−) cells. Total RNA was again collected using Trizol and GenElute LPA carrier per manufacturer instructions. Libraries were prepared using Ribogone kit (Clontech) and the SMARTer Universal Low RNA Kit

(Clontech) and sequenced on the Illumina Hiseq 2500 with post-analysis performed as above using the zebrafish genome.

CRISPR/Cas9 experiment

Cas9 mRNA was produced by *in vitro* transcription from a pCS2 *Cas9* vector (47) using mMESSAGE mMACHINE SP6 kit (Invitrogen). gRNAs were generated following established methods (48). *Sox10* target sequence: GGCCGCGCGCAGGAAACTGG. 600 pg of *Cas9* mRNA and 25 pg of gRNA were injected into embryo of the AB strain. Post microinjection, embryos were raised in E3 medium at 28.5°C. The T7E1 assay was performed as reported (49). Briefly, genomic DNA was extracted from 2-day old embryos using the hotSHOT method {Truett:2000tp}. A fragment of 434 bp was amplified from genomic DNA using the following primers: GAAGTCCGACGAGGAAGAT and CTTGACTGAGTAAATAGTGCGT. The PCR amplicons were then purified on a 1% agarose gel. 200 ng of purified DNA were denatured at 95°C for 5 minutes and slowly reannealed prior to digestion with 10 units of T7E1 enzyme (NEB) for 1 h at 37°C. The digestion product was finally run on a 2.5% agarose gel.

CRISPR/Cas9 tumor-free survival curves

In order to inactivate *sox10* specifically in the melanocytes of our zebrafish melanoma model, the MiniCoopR vector was engineered to express Cas9 under the control of the melanocyte-specific *mitfa* promoter and a gRNA efficiently mutating *sox10*, described above, off a *U6* promoter. A gRNA against *p53* was used as a negative control (50). The two vectors were injected into one-cell stage, Tg(*mitfa:BRAF*^{V600E}), *p53*$^{-/-}$, *mitfa*$^{-/-}$ embryos, and tumor formation was monitored.

To sequence genomic DNA from tumors, tumor tissue was dissected carefully, digested in buffer with proteinase K (51), and after inactivation of proteinase K, PCR was performed using the primers described above as for the T7E1 reaction. PCR fragments were cloned using TopoTA cloning per manufacturer instructions, and colony PCR performed on resulting individual clones and submitted for Sanger sequencing for Figure S9C. For next generation sequencing in Figure S9D, nested PCR (primer sequences TGAACGGGTACGACTGGACGCT and TGTTGTAGCAGTGCGTTTA, yielding a 238 bp amplicon) was performed on the initially amplified genomic locus to bring the amplicon ends closer to the CRISPR target sequence to allow for coverage using a MiSeq-based 150 bp paired-end Illumina run with pooled and barcoded samples at the MGH DNA Core. Both Sanger and compiled next generation sequences were aligned to the wild type locus using Lasergene Seqman to identify changes at the CRISPR target sequence. Wild type *sox10* reads, although potentially from non-targeted *sox10* loci in melanocytes and not necessarily from other tissue types in the sample, were excluded from the calculations of fractions of allele types. If included, these would only increase the fraction of active *sox10* alleles and would further favor our interpretation of the results. When determining the fractions of alleles in panel D, each tumor was weighted equally to avoid skewing from more reads from a given tumor.

Injection of *sox10* mRNA

The *sox10* cDNA from zebrafish was cloned into pENTR/D-TOPO and transferred into pCSDest (52) using Gateway cloning, all per manufacturer instructions. mRNA was generated using SP6 mMessage mMachine Kit (Ambion) per manufacturer, and 1 nanoliter mRNA mix injected into single cell *p53/BRAF/crestin:EGFP* embryos at multiple concentrations with 20 pg being the highest tolerated dose without significant toxicity. Embryos were imaged on a Zeiss Discovery V.8 Stereoscope, and scored for EGFP expression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

CKK and research reported here is supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number K08 AR061071. CM received support from an EMBO long-term fellowship, an HFSP long-term fellowship, and an SNSF advanced postdoctoral fellowship. LIZ is supported by R01 CA103846 and the Ellison Foundation, and he is a Howard Hughes Medical Institute Investigator. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank the MGH DNA Core for providing technical assistance with deep sequencing. We thank Julien Ablain for critical review of the manuscript, Craig Ceol for the gift of the zebrafish *mitfa* middle entry clone, and Gabe Musso for helpful discussions on microarray data and GSEA. All microarray, Chip-Seq, and RNA-Seq will be deposited in GEO.

LIZ is a founder and stock holder of Fate, Inc. and Scholar Rock.

References

1. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer. 1953; 6:963–968. [PubMed: 13094644]
2. Mort, RL.; Jackson, IJ.; Patton, EE. The melanocyte lineage in development and disease. nrs.harvard.edu
3. T. C. G. A. R. Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2015; 511:543–550.
4. Lo JA, Fisher DE. The melanoma revolution: from UV carcinogenesis to a new era in therapeutics. Science. 2014; 346:945–949. [PubMed: 25414302]
5. Patton EE, et al. BRAF Mutations Are Sufficient to Promote Nevi Formation and Cooperate with p53 in the Genesis of Melanoma. Current Biology. 2005; 15:249–254. [PubMed: 15694309]
6. Luo R, An M, Arduini BL, Henion PD. Specific pan-neural crest expression of zebrafish Crestin throughout embryonic development. Dev Dyn. 2001; 220:169–174. [PubMed: 11169850]
7. Rubinstein AL, Lee D, Luo R, Henion PD, Halpern ME. Genes dependent on zebrafish cyclops function identified by AFLP differential gene expression screen. Genesis. 2000; 26:86–97. [PubMed: 10660676]
8. White RM, et al. DHODH modulates transcriptional elongation in the neural crest and melanoma. Nature. 2011; 471:518–522. [PubMed: 21430780]
9. Mosimann C, et al. Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. Development. 2011; 138:169–177. [PubMed: 21138979]
10. Kong Y, et al. Neural Crest Development and Craniofacial Morphogenesis Is Coordinatedby Nitric Oxide and Histone Acetylation. Chem Biol. 2014:1–14. [PubMed: 24439204]
11. Grzmil M, et al. The INT6 Cancer Gene and MEK Signaling Pathways Converge during Zebrafish Development. PLoS ONE. 2007; 2:e959. [PubMed: 17895999]

12. Hildemann WH. Tissue Transplantation Immunity in Goldfish. Immunology. 1958; 1:46. [PubMed: 13513141]
13. Vogel AM, Gerster T. Promoter activity of the zebrafish bhikhari retroelement requires an intact activin signaling pathway. Mech Dev. 1999; 85:133–146. [PubMed: 10415354]
14. Sandelin A, Alkema W, Engström P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res. 2004; 32:D91–4. [PubMed: 14681366]
15. Ceol CJ, et al. The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. Nature. 2011; 471:513. [PubMed: 21430779]
16. Uhlen M, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics. 2005; 4:1920–1932. [PubMed: 16127175]
17. Bosserhoff AK. Melanoma inhibitory activity (MIA): an important molecule in melanoma development and progression. Pigment Cell Res. 2005; 0:051012082332003.
18. Weinlich G, et al. Metallothionein – overexpression as a highly significant prognostic factor in melanoma: a prospective study on 1270 patients. British Journal of Cancer. 2006; 94:835–841. [PubMed: 16508630]
19. Kelsh RN. Sorting out Sox10 functions in neural crest development. Bioessays. 2006; 28:788–798. [PubMed: 16927299]
20. Kim YJ, et al. Generation of Multipotent Induced Neural Crest by Direct Reprogramming of Human Postnatal Fibroblasts with a Single Transcription Factor. Cell Stem Cell. 201410.1016/j.stem.2014.07.013
21. Cronin JC, et al. SOX10 Ablation Arrests Cell Cycle, Induces Senescence, and Suppresses Melanomagenesis. Cancer Res. 2013; 73:5709–5718. [PubMed: 23913827]
22. Graf SA, Busch C, Bosserhoff AK, Besch R, Berking C. SOX10 Promotes Melanoma Cell Invasion by Regulating Melanoma Inhibitory Activity (MIA). J Investig Dermatol. 201410.1038/jid.2014.128
23. Shakhova O, et al. Sox10 promotes the formation and maintenance of giant congenital naevi and melanoma. Nature Cell Biology. 2012; 14:882–890. [PubMed: 22772081]
24. Rada-Iglesias A, et al. Epigenomic Annotation of Enhancers Predicts Transcriptional Regulators of Human Neural Crest. Cell Stem Cell. 201210.1016/j.stem.2012.07.006
25. Olesnicky E, Hernandez-Lagunas L, Artinger KB. Prdm1a regulates sox10 and islet1 in the development of neural crest and Rohon-Beard sensory neurons. Genesis. 201010.1002/dvg.20673
26. Hnisz D, et al. Super-Enhancers in the Control of Cell Identity and Disease. Cell. 2013; 155:934–947. [PubMed: 24119843]
27. Parker SCJ, et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proceedings of the National Academy of Sciences. 2013; 110:17921–17926.
28. Whyte WA, et al. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell. 2013; 153:307–319. [PubMed: 23582322]
29. Lovén J, et al. Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. Cell. 2013; 153:320–334. [PubMed: 23582323]
30. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Meth. 2013; 10:1213–1218.
31. Barretina J, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483:603–607. [PubMed: 22460905]
32. Rada-Iglesias A, et al. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011; 470:279–283. [PubMed: 21160473]
33. Youssef KK, et al. Adult interfollicular tumour-initiating cells are reprogrammed into an embryonic hair follicle progenitor-like fate during basal cell carcinoma initiation. Nature Cell Biology. 2012; 14:1282–1294. [PubMed: 23178882]
34. Rönnstrand L, Phung B. Enhanced SOX10 and KIT expression in cutaneous melanoma. Med Oncol. 2013; 30:648. [PubMed: 23801280]

35. Yen J, et al. The genetic heterogeneity and mutational burden of engineered melanomas in zebrafish models. Genome Biol. 2013; 14:R113. [PubMed: 24148783]
36. Kwan KM, et al. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. Dev Dyn. 2007; 236:3088–3099. [PubMed: 17937395]
37. Tamplin OJ, Cox BJ, Rossant J. Integrated microarray and ChIP analysis identifies multiple Foxa2 dependent target genes in the notochord. Developmental Biology. 2011; 360:415–425. [PubMed: 22008794]
38. Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat Protoc. 2008; 3:59–69. [PubMed: 18193022]
39. Lee TI, Johnstone SE, Young RA. Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat Protoc. 2006; 1:729–748. [PubMed: 17406303]
40. Robinson JT, et al. Integrative genomics viewer. Nat Biotechnol. 2011; 29:24–26. [PubMed: 21221095]
41. Kent WJ, et al. The Human Genome Browser at UCSC. Genome Research. 2002; 12:996–1006. [PubMed: 12045153]
42. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Meth. 2012; 9:357–359.
43. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]
44. Heilmann S, et al. A Quantitative System for Studying Metastasis Using Transparent Zebrafish. Cancer Res. 2015; 75:4272–4282. [PubMed: 26282170]
45. Howe DG, et al. ZFIN, the Zebrafish Model Organism Database: increased support for mutants and transgenics. Nucleic Acids Res. 2013; 41:D854–60. [PubMed: 23074187]
46. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005; 102:15545–15550. [PubMed: 16199517]
47. Jao LE, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proceedings of the National Academy of Sciences. 2013; 110:13904–13909.
48. Gagnon JA, et al. Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. PLoS ONE. 2014; 9:e98186. [PubMed: 24873830]
49. Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Research. 2009; 19:1279–1288. [PubMed: 19470664]
50. Ablain J, Durand EM, Yang S, Zhou Y, Zon LI. A CRISPR/Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish. Dev Cell. 201510.1016/j.devcel.2015.01.032
51. Wienholds E, et al. Efficient target-selected mutagenesis in zebrafish. Genome Research. 2003; 13:2700–2707. [PubMed: 14613981]
52. Villefranc JA, Amigo J, Lawson ND. Gateway compatible vectors for analysis of gene function in the zebrafish. Dev Dyn. 2007; 236:3077–3087. [PubMed: 17948311]

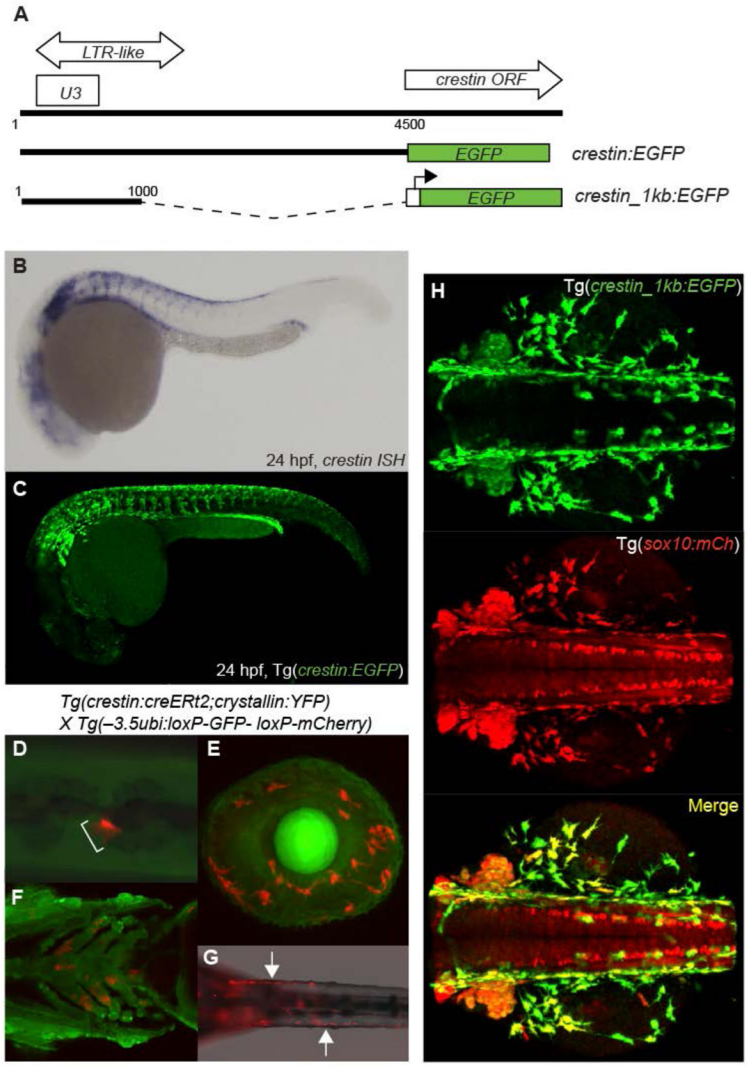


Figure 1. The *crestin* promoter/enhancer drives neural crest-specific gene expression

(A) Prototypical *crestin* retrotransposon locus with predicted ORF, LTR-like, and U3-like promoter regions. Locations of 4.5 kb and 1 kb segments used for *crestin:EGFP* constructs (white box/promoter arrow indicate β-globin gene minimal promoter). **(B)** Endogenous expression pattern of *crestin* transcript by ISH (purple staining) at 24 hpf marks developing and migrating neural crest cells. **(C)** This expression pattern (green) is recapitulated by a stable *Tg(crestin:EGFP)* embryo at 24 hpf. **(D–G)** Genetic lineage tracing of cells that express *crestin* (*Tg(crestin:creERt2;crystallin:YFP)* X *Tg(–3.5ubi:loxP-GFP-loxP-mCherry)*) marks multiple neural crest lineages (red cells) including melanocytes (bracket) on the dorsum **(D)** and the eye **(E)** (72 hpf), **(F)** jaw cartilage (ventral view, 5 dpf), and **(G)** glial cells of the lateral line (arrows, dorsal view posterior to the yolk, 72 hpf). **(H)** *Tg(crestin:EGFP)* expression overlaps significantly with a *sox10:mCh* transgene (confocal image, dorsal view over yolk, 24 hpf).

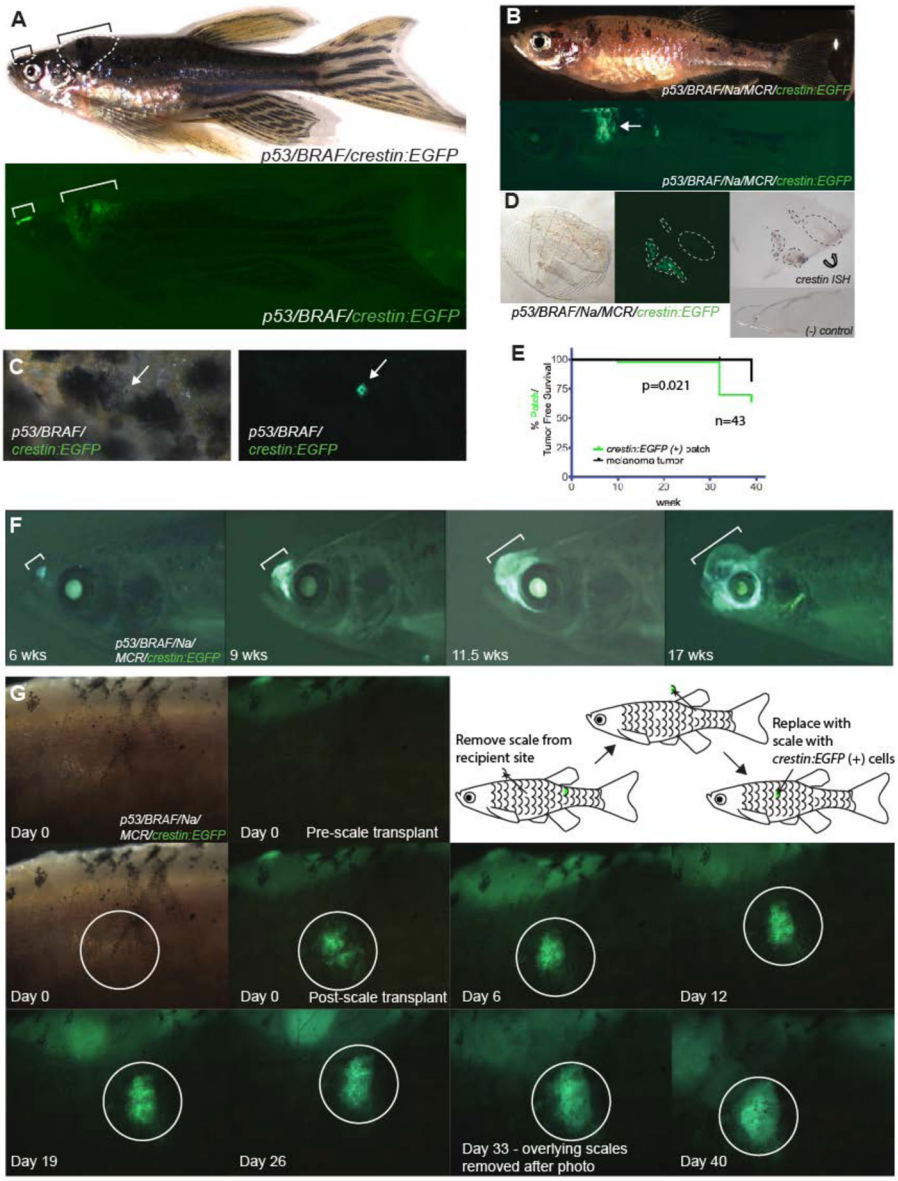


Figure 2. *Tg(crestin:EGFP)* specifically marks melanoma tumors and precursor lesions

(A) Spontaneously arising tumors in *p53/BRAF/crestin:EGFP* zebrafish express *EGFP* (brackets) whereas the remainder of the animal is negative. **(B–C)** Scales expressing *crestin:EGFP* from precursor, non-raised regions (arrow, lower panel) were plucked, photographed **(C,** left and middle panel), and subjected to ISH for *crestin* transcript **(C,** right panel). Note the concordance of *EGFP* (green) and *crestin* transcript (purple, dotted outlines, scales curl during ISH procedure, indicated by the curved arrow, observed in 5/5 scales). **(C,** lower right**)** *crestin:EGFP* negative scales are negative for *crestin* ISH staining (observed in 7/7 tested scales). **(D)** Example of a single *crestin:EGFP* (+) cell in *p53*/*BRAF* background. **(E)** Cohorts of *p53/BRAF/crestin:EGFP* zebrafish were tracked over time for the appearance of *crestin:EGFP* (+) patches and tumors, with *crestin:EGFP* (+) cells/patches (green line) identifiable prior to raised melanoma tumors (black line). **(F)** Example of an *EGFP* (+)

preclinical patch tracked over time (6, 9, 11.5, 17 weeks) as it expands into a clinically apparent melanoma tumor. **(G)** Scale autotransplant and expansion of *crestin:EGFP* (+) patch of cells. At Day 0, the recipient site is free of *crestin:EGFP* (+) cells pre-scale transplant, but immediately following transplant of a single scale (post-scale transplant), the patch of *EGFP* (+) cells is apparent (white circle). This patch expands outward, and even upon removal of the original transplanted scale after the day 33 photograph, *EGFP* (+) cells remain in place and continue to expand. Same magnification and size of white circle in each image.

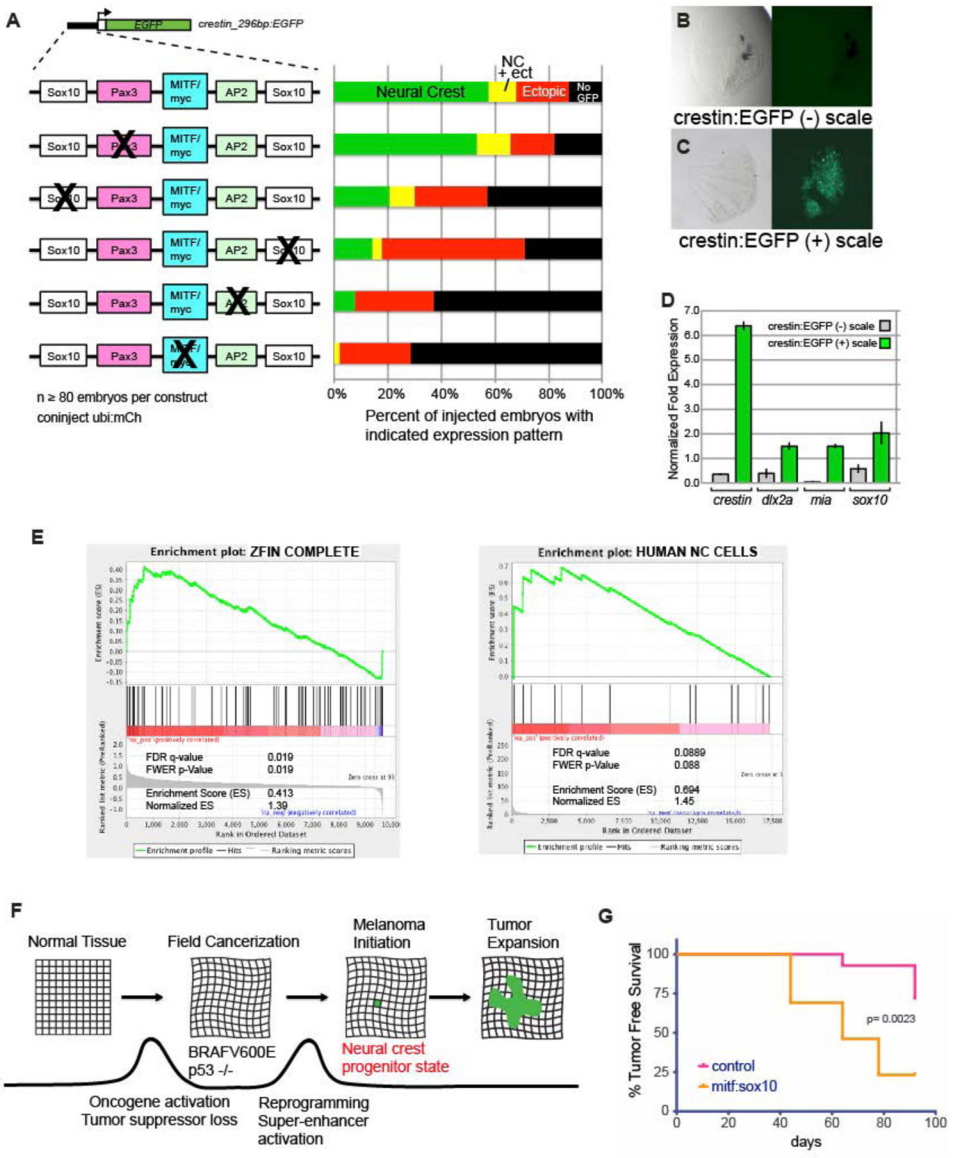
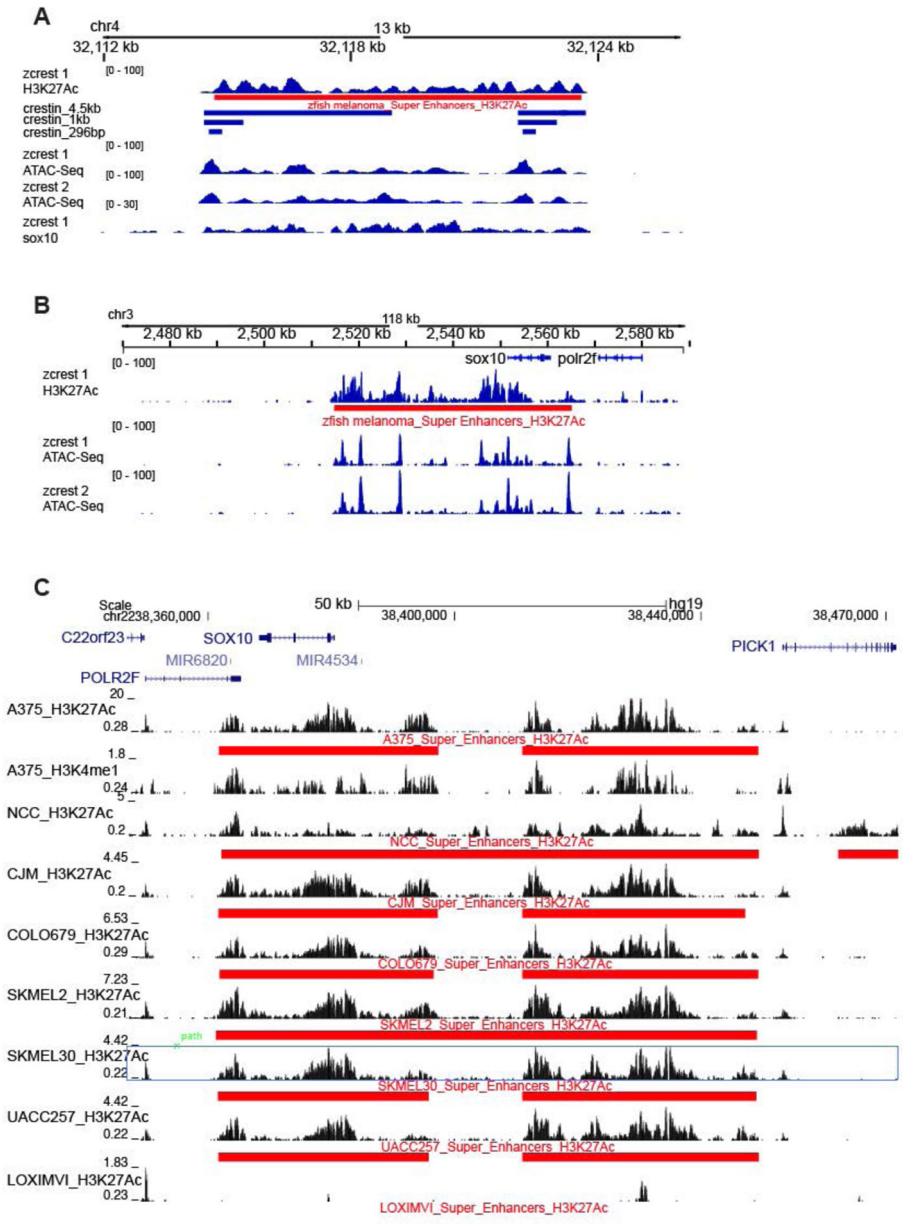


Figure 3. Reemergence of neural crest progenitor identity in melanoma initiation

(A) Mutation of key neural crest transcription factor binding sites in the 296 bp *crestin* element, including *sox10, tfap2,* and an E-box for *myc* or *mitf,* substantially reduces neural crest *EGFP* expression at 24 hpf, whereas mutation of the predicted *pax3* site does not alter expression. Coinjection of a ubiquitous *ubi:mCh* transgene confirmed successful injection for the >80 independently injected F0 embryos analyzed for each construct. Scales from *p53/BRAF/Na/MiniCoopR/crestin:EGFP* adult zebrafish with (**B**) and without (**C**) *EGFP* (+) cells were collected, and total RNA isolated for microarray analysis. **(D)** Quantitative RT-PCR of *crestin:EGFP* (+) versus (−) scales reveals enrichment of neural crest *(crestin, dlx2a, sox10)* and melanoma marker expression (*crestin, mia, sox10*). **(E)** GSEA analysis shows a positive association between *crestin:EGFP* (+) patch-enriched genes and neural crest-expressed genes in zebrafish (left panel) and in human ES-derived neural crest cells (right panel). **(F)** Model for the importance of reemergence of neural crest progenitor (NCP) state

through SE activation as an essential step in melanoma initiation. The acquisition of genetic lesions in normal tissue leads to oncogene activation (*BRAF*^{V600E}) and tumor suppressor loss (*p53−/−*) and represents an initial barrier that generates a cancerized field from which rare clones (green) overcome the additional barrier of achieving a NCP state to initiate melanoma formation and then tumor expansion. Favoring reemergence of the neural crest progenitor state would then increase melanoma formation, and strengthening this barrier to inhibit adoption of the *crestin* (+) NCP state would block melanoma initiation **(G)** Misexpression of the NCP transcription factor *sox10* accelerates melanoma onset compared to controls in *p53/BRAF/Na* zebrafish rescued with the *MiniCoopR* construct.



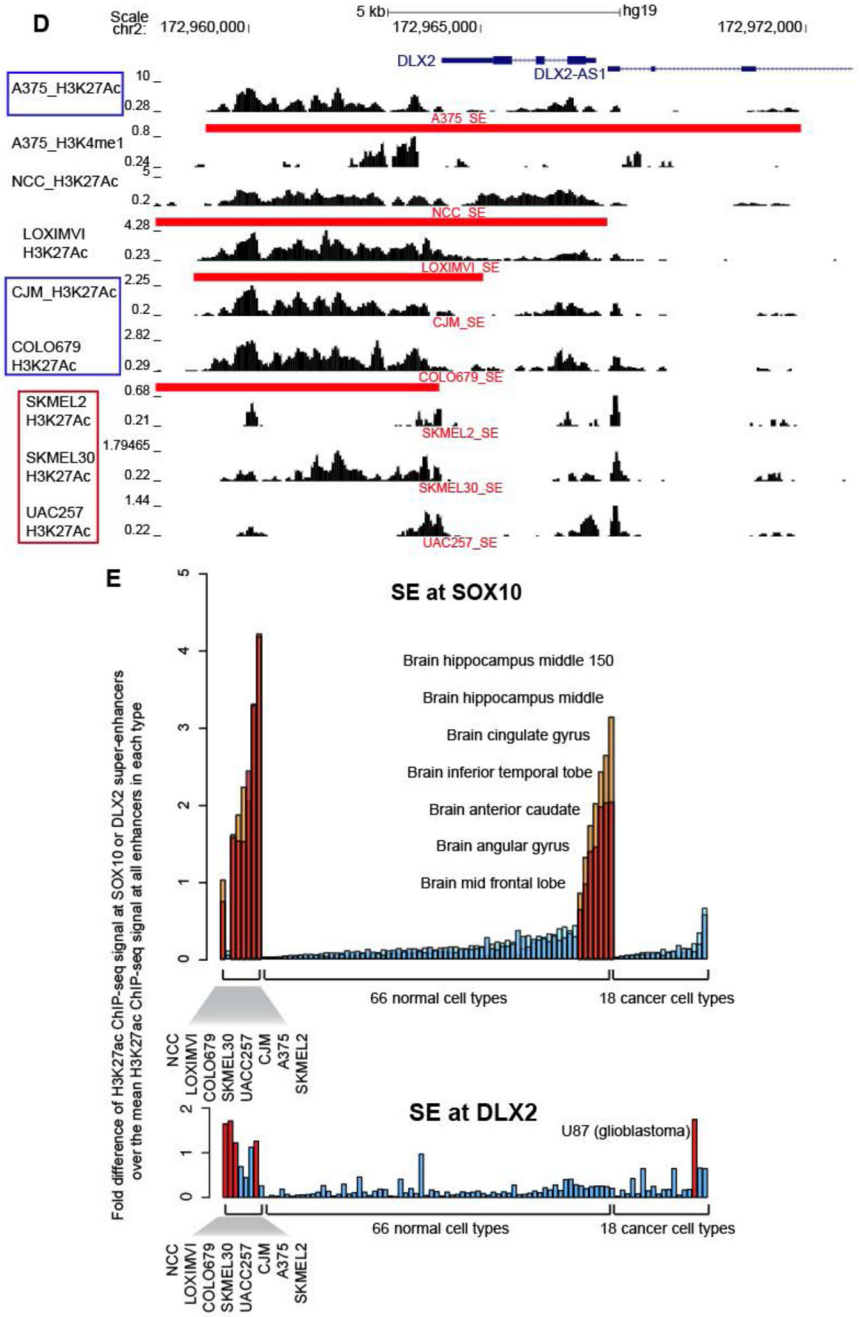


Figure 4. A super-enhancer (SE) signature in zebrafish and human melanoma

(A) ChIP-Seq for the H3K27Ac histone mark (top row) in a *crestin:EGFP (+)* zebrafish melanoma cell line (zcrest 1) reveals enriched peaks, identified as a SE (red bar), at a representative *crestin* locus. Sequences of *crestin_4.5kb, crestin_1kb,* and *crestin_296bp* shown with blue horizontal bars. ATAC-Seq on two zebrafish melanoma lines (zcrest 1 and zcrest 2) identifies open chromatin coincident with the H3K27Ac marks at *crestin* loci. ChIP-Seq for Sox10 shows enrichment across the *crestin* locus (bottom row) in zcrest 1 cells. **(B)** ChIP-Seq for the H3K27Ac histone mark (top row) on the zcrest 1 line identifies

robust enrichment and a SE at *sox10* (red bar). ATAC-Seq identifies corresponding regions of open chromatin (rows 2 and 3). **(C)** ChIP-Seq for the H3K27Ac mark on multiple *SOX10*-expressing melanoma lines (A375, CJM, COLO679, SKMEL2, SKMEL30, UACC257) and a rare *SOX10*-negative melanoma line (LOXIMVI). Robust peaks corresponding to SE's (red bars) are identified in all lines, except the *SOX10*-negative LOXIMVI line. Published H3K27Ac ChIP-Seq data from ES-derived human neural crest cells (NCC's) reveals a similar SE pattern. ChIP-Seq for H3K4Me1, an enhancer mark, on a representative melanoma line, A375 (row 2), identifies regions corresponding to the H3K27Ac marks. **(D)** H3K27Ac signal is robust at the *DLX2* locus in melanoma cell lines not expressing the melanocyte differentiation genes *TYR* and *DCT* (blue box), in human NCC's, and in the *SOX10*-negative LOVIMVI melanoma line. Human genomic track images generated at http://genome.ucsc.edu. **(E)** High relative H3K27ac signal at *SOX10* (upper panel) and *DLX2* (lower panel) identifies SE's (presence = red/orange bar, absence = blue bars) and is largely enriched in melanomas and human NCC's compared to 66 normal and 18 cancer cell types.